



Beyond the promise: implementing ethical AI

Ray Eitel-Porter¹

Received: 31 August 2020 / Revised: 31 August 2020 / Accepted: 3 September 2020
© The Author(s) 2020

Abstract

Artificial Intelligence (AI) applications can and do have unintended negative consequences for businesses if not implemented with care. Specifically, faulty or biased AI applications risk compliance and governance breaches and damage to the corporate brand. These issues commonly arise from a number of pitfalls associated with AI development, which include rushed development, a lack of technical understanding, and improper quality assurance, among other factors. To mitigate these risks, a growing number of organisations are working on ethical AI principles and frameworks. However, ethical AI principles alone are not sufficient for ensuring responsible AI use in enterprises. Businesses also require strong, mandated governance controls including tools for managing processes and creating associated audit trails to enforce their principles. Businesses that implement strong governance frameworks, overseen by an ethics board and strengthened with appropriate training, will reduce the risks associated with AI. When applied to AI modelling, the governance will also make it easier for businesses to bring their AI deployments to scale.

Keywords AI · Ethics · Responsible AI · Ethical AI · AI principles · Data governance

1 Introduction

Ethical AI, also known as responsible AI, is the practice of using AI with good intention to empower employees and businesses, and fairly impact customers and society. Responsible AI enables companies to engender trust and scale AI with confidence.

Around the world, a growing number of organisations are working on ethical AI principles and frameworks. These include academia-led programmes, such as The Institute for Ethical AI and Machine Learning, trade union-led schemes, such as UNI Global Union, and business-led initiatives, such as Microsoft's responsible AI guidelines [1–3].

These and other such initiatives reflect a growing acknowledgement that AI applications can and do have unintended negative consequences if not implemented carefully. More broadly, ethical AI is part of a wider responsible business agenda, whereby organisations are increasingly prioritising good governance and a respect for the societal and environmental concerns of customers [4].

Although ethical principles are a necessary precondition for responsible AI, they are not sufficient. Ethical standards only have value when put into practice. In this paper, I argue that responsible AI also requires strong, mandated governance controls including tools for managing processes and creating associated audit trails. I also argue that good governance helps businesses scale their AI tools and extract full value from their AI applications and services.

For the purposes of this paper, I focus solely on the trust, fairness, and privacy elements of AI deployments. Although related, this paper does not examine issues concerning data security.

The topic of AI ethics and governance is a timely one. According to Gartner, a research and advisory company, by 2022, 85% of AI projects could deliver erroneous outcomes due to bias in data, algorithms, or the teams responsible for managing them [5]. Meanwhile, research by Accenture suggests that while 63% of leaders believe it is crucial to monitor AI systems, most are unsure how to. Sixty percent require a human override of an AI system at least once a month [6]. Challenges, such as these, need to be addressed directly if AI is to deliver maximum value with minimal risk. Accenture's Applied Intelligence practice, which incorporates a dedicated capability for Responsible AI, is specifically focused

✉ Ray Eitel-Porter
ray.eitel-porter@accenture.com

¹ European Lead for Responsible AI at Accenture, London, UK

on the application of AI, rather than research or Corporate Social Responsibility [7].

2 Common risks and pitfalls associated with corporate AI deployments

Applied AI can be used with the intent of causing harm, such as through autonomous weapons and social engineering [8]. In such cases, ethical norms are discarded, and therefore such applications are out of the scope of the current discussion.

The challenge faced by most businesses is around limiting those consequences of AI that are *unintentionally* negative. Broadly speaking, unintended consequences arise when AI is deployed without sufficiently robust governance and compliance efforts. They fall into three categories:

1. **Compliance and governance.** The risk of breaching regulations around, for example, employment, data privacy, financial services and health and safety. For instance, biases in training data can cause recruitment apps to favour one gender over another, which breaches anti-discrimination laws. For this very reason and according to Reuters, Amazon's machine learning team halted the development of its talent evaluation app in 2015, having been trained mostly on data derived from the professional resumes of men, the algorithm taught itself to favour male job applicants over female [9].
2. **Brand damage.** The risk of breaching social norms and taboos. If an AI threatens to cause outrage or offence, it can damage the reputation of the company that launched it. One example is Microsoft's Tay chatbot, which was programmed to learn from conversations with Twitter users. Some of these users quickly began to use and feed inflammatory and racist language, which was learnt and repeated by the chatbot, and Microsoft shut it down the next day [10]. That same year, Amazon came under fire because its same day delivery service was initially only offered to predominantly white, affluent areas of the US. The apparent reason was that based on the available data about concentration of Prime™ members and proximity to warehouses, Amazon's algorithm calculated that the service would not be profitable in areas with large ethnic minority populations [11].
3. **Third-party transparency.** The risk stemming from unexplainable, black box AI tools from third-party providers. At Accenture, we have seen a case where an organisation used third-party algorithms as black boxes with no detailed understanding of how they work. Independent review has subsequently revealed biases that could have caused brand and compliance issues for this organisation.

Often, the more innovative an organisation is, the greater the risks it takes. But to reduce the possibility of falling foul of these risks, businesses deploying AI need to be aware of common pitfalls. These include:

- **Rushed development** may result in AI that appears to function well but which contains underlying problems because corners have been cut in an effort to meet deadlines.
- **A lack of technical understanding** of the proper and responsible use of AI hampers development from the outset. Given the lack of data science skills in the IT space in general and the complexity of AI programming, this is a common challenge [12].
- **Improper quality assurance** due to constraints in understanding fails to pick up issues. This can happen when quality assurance is conducted by the application developer rather than by an independent examiner, or when the individual conducting the quality assurance lacks the requisite technical understanding to identify potential issues. AI quality assurance also requires new skills, which some organisations may struggle to find [13]. Indeed, some educational establishments are mooted the need for specific courses or modules devoted to training data scientists in techniques for ethical assurance.
- **Use of AI outside of original context** can cause unexpected results. This is because the AI programmers may not have accounted for variables outside of their original focus.
- **Improper combination of data** may expose companies to risk. Firms should collect no more data than is required by the AI model. The more data a firm gathers, the higher the risk of non-compliance with privacy regulations, such as the General Data Protection Regulation (GDPR).
- **Reluctance by employees to raise concerns** means issues go unreported. Companies should put in place clear mechanisms for employees to communicate concerns around AI models to their superiors and to ensure these concerns are acted on without repercussions for those employees.

Within most enterprises, we consider there are three areas where questions around ethical AI are most important. The first is technical and relates to how the AI models perform. Are the models both accurate and fair, or are they biased in some negative way? Noting that bias can result from faulty design, data selection or calculation method.

The second relates to the organisation and its workers. The key considerations here are whether an AI model can cause job losses, or adversely impact the day-to-day experience of employees. A case currently subject to trial in the courts illustrates this point. A group of Uber drivers

in the UK has launched a legal action seeking algorithmic accountability and transparency. The plaintiffs allege that Uber failed to provide them with sufficient information on the automated decision-making and profiling that takes places in the Uber Driver App, and that such profiling could mean algorithmic management of drivers, something that allegedly goes against Uber's traditional stance that drivers are self-employed and are, therefore, not subject to management control. Such lack of information is something that, drivers argue, is in breach of GDPR. This is relevant because Uber's algorithms essentially determine the earning potential of a driver based on how the platform assigns jobs [14].

The third element is brand reputation, as mentioned above. This is increasingly important as companies grow to rely on AI as a major touch point with customers (according to Gartner, the use of AI by enterprises grew 270% in the 4 years up to 2019) [15]. Just as companies invest time and effort training call centre workers to reflect their brand and treat customers with high levels of service, so too should organisations spend time training their AI channels in brand voice and expected service levels.

3 The key principles of responsible AI

The way in which organisations can address the risks of AI through a principles-based approach is now well understood. Over the past two to three years, we have seen organisations focus on defining core principles. While different organisations/initiatives will have slightly different takes on what their list of responsible AI principles should include, they generally share the same five pillars:

1. **Fairness.** Are there factors influencing model outcomes that should not be? Is there an expectation of similar outcomes for different subgroups, and is this delivered?
2. **Accountability.** What is the chain of command for dealing with a potentially biased/erroneous outcome?
3. **Transparency.** Do we understand how the model works?
4. **Explainability.** Can we explain, in non-technical language, why an output was arrived at?
5. **Privacy.** Does the model ensure against inferences that breach privacy?

If followed, these principles should help ensure an AI model is responsible. However, the key phrase is 'if followed'. On their own, principles offer only a sign post. To be effective, they need to be applied rigorously.

4 Ethical AI governance in business

The effective application of responsible AI by businesses requires an effective cross-functional governance structure, training at all levels, and appropriate tools to support the governance framework.

These elements should come together in an end-to-end methodology that reinforces the AI principles at every stage of the journey from proof of concept to production.

Step 1: Founding principles.

The governance journey starts with an organisation's chosen ethical principles. As mentioned, there are a large number of frameworks already in existence that companies can use as a starting point. These include everything from Asimov's Three Laws of Robotics to the more recent (and more detailed) IEEE The General Principles of Ethically Aligned Design, the UK House of Lords Select Committee: Five Core Principles to Keep AI Ethical, and the European Commission's High Level Expert Group on AI Ethics Guidelines for Trustworthy AI to name a few [16–19].

In this process, it is important that companies arrive at a final set of principles that suit their unique circumstances and their brand values. For instance, a company in the chemicals manufacturing industry may have identified its core business priority as safety, and used this value to define its operational processes, customer messaging, value propositions and much else besides. As a new manifestation of the brand, safety should also be an integral element of its AI principles.

Setting out, companies should, therefore, ask themselves questions, such as what are my company's core values? How might these values be reflected in my AI? What are the most profound ethical issues facing my company? And what are the strategic goals of my company [20]? Working back from the answers to these questions, companies can map out their responsible AI principles.

Step 2: Establish an ethics board.

A second foundational element for implementing responsible AI within organisations is to put in place an ethics board. An ethics board comprises subject matter experts and ethics experts and its function is to provide advice and approval for broad AI strategies and specific use cases. The scope of the ethics board will vary from company to company according to organisational size and objectives.

Step 3: Establish a governance structure.

The next step is to identify a robust governance process to clarify how decisions around responsible AI are made and documented. For companies in sectors that already require governance structures and cultures, such as financial services, a good approach is to adapt the existing structure to the requirements of responsible AI. This is

because employees will already be familiar with these processes and will likely feel more comfortable using them to report concerns. Reusing existing structures also removes duplication of effort and helps prevent two competing sources of governance within a single enterprise.

When identifying processes and apportioning responsibility for decision-making, it is important to provide clarity over who is expected to make the trade-off decisions that will inevitably come with AI. These are the areas where what is desirable from a technical standpoint may not be desirable from an ethical standpoint. For instance, an insurance company may have to trade off the accuracy of policy pricing models to ensure compliance with regulations—for example, car insurance algorithms programmed to ignore the wealth of data that shows women are less likely than men to be involved in accidents [21].

This is a fairly black and white case, but there are others where the trade-offs will be less obvious and more thought will be required as to the best course of action to take as an organisation. As part of their governance structures, businesses need to know who will make the final decision on these trade-offs and ensure that tooling is available to capture who made the decision, when, and why, in case the decision is later challenged by a customer, regulator or other stakeholder.

Step 5: Implement company-wide training.

Training is essential to ensure that corporate AI principles and governance structures are understood and acted on by all levels of the organisation. Businesses should commence with a GAP analysis to identify training priorities and then proceed with training tracks around the policies and procedures that have been developed by the leadership, impact and risk assessment tools, and the basics of AI and machine learning technology.

Training will vary according to seniority and role. Senior leaders, for instance, should set the tone by learning why responsible AI matters and then embodying this learning. Middle management needs to understand how issues of responsible AI may affect their projects and what the governance procedures are. And front-line data scientists need to understand that controlling for fairness and bias in AI is complex and requires specialist skills.

Step 6: Challenge the governance structure.

Stress tests are important to ensure that a company's responsible AI governance procedures are robust and resilient. Independent challenge and oversight, therefore, needs to be embodied in governance structures.

One approach is to establish 'Red Teams' and 'Fire Wardens'. The concept of Red Teaming is borrowed from the world of cybersecurity and relates to the use of 'white hat' hackers to test enterprise defences [22]. Applied to responsible AI, Red Teams constitute data scientists

charged with reviewing algorithms and outcomes for signs of bias or the risk of unintended consequences.

The Fire Warden approach considers instances of bias or unethical uses of AI as fires. With real fires, fire wardens are trained to raise the alarm and then carry out the required safety actions. So too, AI Fire Wardens can be trained to follow a clear set of protocols to escalate issues relating to AI as they are encountered.

Step 7: Establish metrics.

Alongside these stress tests, organisations should put in place clear metrics to ensure that AI principles are followed and that they deliver the positive impact required of them. It is likely that without such measures in place, responsible AI principles will be abandoned as business units press ahead with daily priorities. However, organisations should also be wary of overmeasurement as it can slow AI development; it is no accident that the UK's body tasked with advising on ethical standards for AI is called the Centre for Data Ethics and Innovation.

Step 8: Foster dissent and promote diversity.

Dissent is an important element of a robust governance framework. Where there is a breakdown in communication, failure of leadership, or improper risk assessment, employees need to feel comfortable voicing their concern.

It is also important to avoid monocultures within the teams tasked with designing AI. A team consisting of people from diverse backgrounds helps guard against unconscious bias. A 2019 study suggests that this particular area of governance should be addressed as a priority. The study, which was conducted by The AI Now Institute at New York University, found that the AI field, which is overwhelmingly white and male, is at risk of replicating or perpetuating historical biases and power imbalances [23].

Examples of how gender imbalance in data compounds inequality are not hard to come by. Caroline Criado Perez' book *Invisible Women* argues how much of the modern world is built around the needs of men, due in a large part to data bias. To give just one instance, according to Caroline's research, most offices are five degrees too cold for women. This is because the formula was based on the metabolic resting rate of a 40-year-old, 70 kg man, whereas women's metabolisms are slower [24].

5 AI governance applied to model building

Responsible AI governance has a role to play in all stages leading from the modelling of an AI application to its deployment at scale:

1. **Data selection and model building.** Mandate a review and document a sign-off process against a detailed set of structured questions.

These questions relate to every element of the data and the model (i.e. selection of data set(s), design, modelling algorithms, and execution) and focus on areas, such as the system (e.g. is the system's existence responsible? Is the technical approach judicious? Is the model effective and well built?) bias (e.g. are the data used appropriately? Is the model behaviour understood and fair? Is the environment around the tool a likely source of bias?) and transparency (e.g. are the data understood and documented? Is the model explainable? Is the tool sufficiently monitored and measured?).

Businesses can leverage catalogues of questions, explanations, and archetypical answers provided by third-party consultancies and partners to serve as the foundation of the assessment. It is important that these questions are modified to each organisation's situation. Tooling can be used to automate some of this process, sending documents to the independent reviewer for sign-off before the next stage is initiated.

2. **Pre-production.** As the AI model is being readied for production, it is essential to capture appropriate documentation, such as information, on the training data used, the code and answers to assessment questions, and names of the approvers. This information is important to provide an audit trail in case of subsequent issues with the model. It can also be used to enable teams to manage the live model when they have no recourse to the original author.

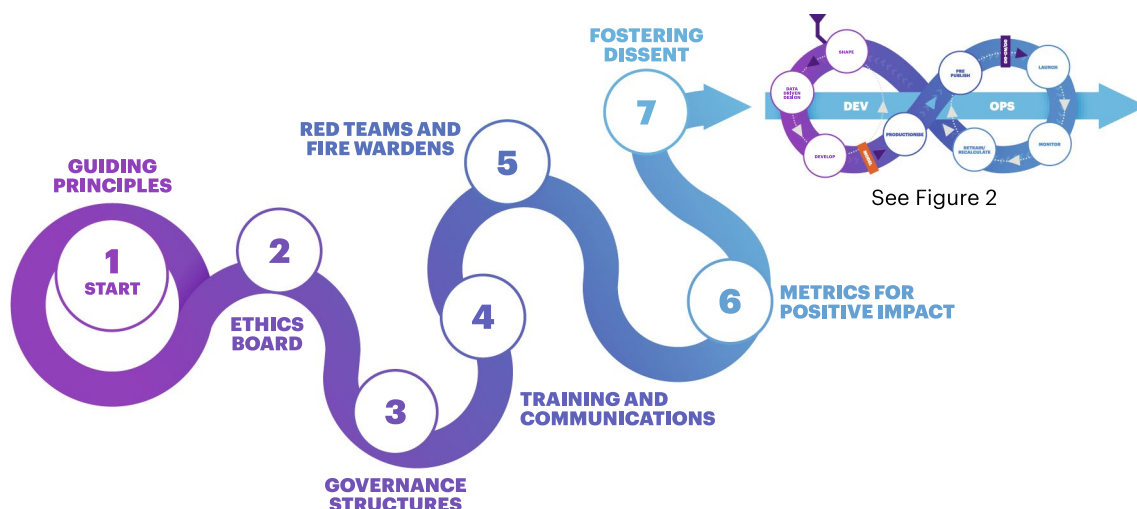
Before a model is put forward for production, it should pass a hard phase gate that confirms the model has met the appropriate checks and independent review. Accenture applies these principles for its internal use of AI models.

3. **Model management.** Responsible AI governance does not stop once a model is launched. Models need to be reviewed continually to ensure that ethical parameters are not breached over time.

Managers can learn from processes currently used to manage model technical performance over time. The automated monitoring of live models is used to ensure that given parameters for the effective functioning of the AI do not move outside of predetermined bounds. These automated systems are often capable of making small adjustments on-the-fly, but their prime function is to alert human data scientists to performance issues so they can be fixed.

Such systems can also be used to detect changes over time that cause bias to emerge. Suppose a hospital deployed a diagnostic tool primed to suit the needs of its local community, which is dominated by young people. If over time the local area changed and large numbers of elderly people were to move in, these elderly patients might miss out on the benefits of the diagnostic tool as the AI is not configured to take account of their needs. An automated alert system would pick up this change in demography as a material issue and flag it to the model maintenance team.

The various ways in which responsible AI governance should influence AI production in organisations can be seen in the following two figures. The first represents the end-to-end process, from principles to scale, while the second focuses on the practical application of governance in the process of model development and deployment (Figs. 1, 2).



Copyright © 2020 Accenture. All rights reserved.

Fig. 1 Scaling the responsible use of AI

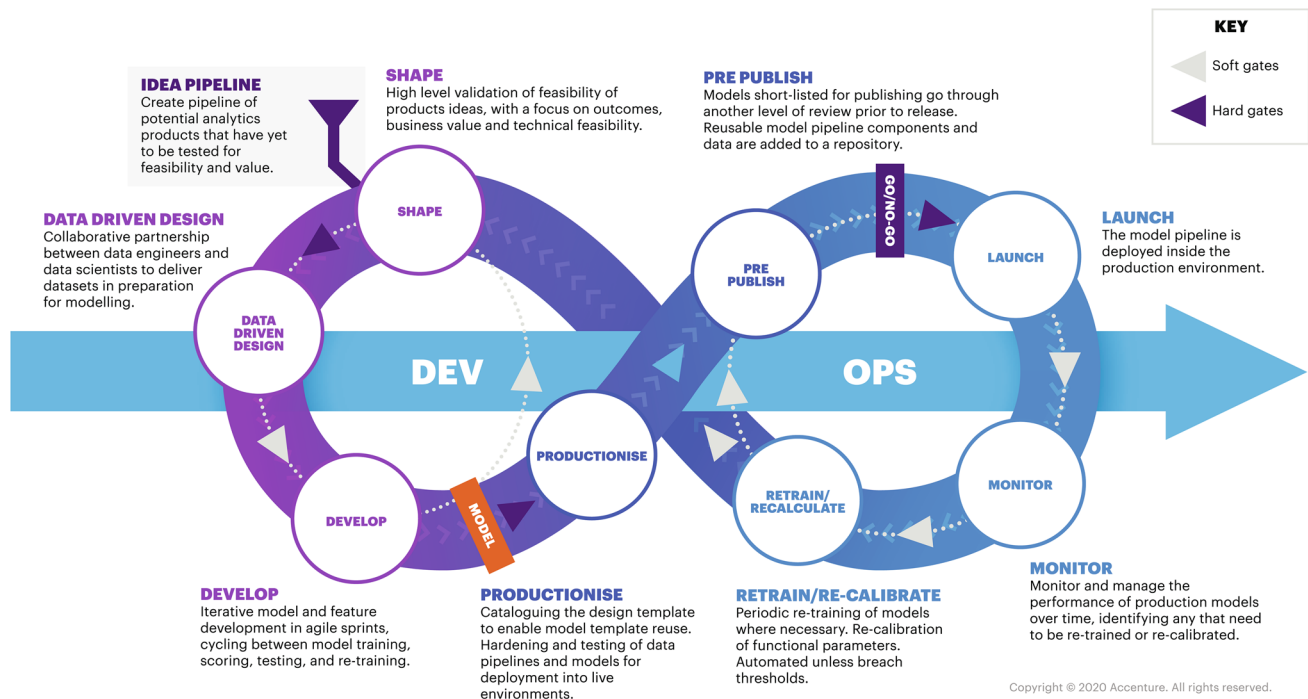


Fig. 2 Detail of continuous AI engineering model

6 AI at scale

To a large extent, the issue of responsible AI governance is linked to that of successfully bringing AI applications to scale. Where companies have only a limited number of AI models, these models can be managed by the original developers and maintained using little more than an Excel sheet. Challenges arise when these models proliferate and are scaled up. New teams come on board to maintain the algorithms and risk losing sight of their core purpose and technical functionality, data sets grow and bias may creep in, generating unexpected complications and affecting performance.

Bringing AI to scale is proving to be a significant stumbling block for many organisations. Research from Accenture has found that 87% of companies are struggling to realise the full value of their AI projects and move beyond proof of concept to production, because there is no clear path to deployment [25]. The research found that to scale effectively, organisations need to have a clear AI strategy, diverse teams and ethical frameworks built into their AI, among other things.

At this point, controls for responsible AI and controls for high-performing AI converge. Industrial scale AI requires good governance, documentation, and tooling to be both ethical and effective. Installing a rigorous governance programme for the one, helps achieve the other. This is important because companies that are strategically scaling AI

report nearly three times the return from AI investments compared to companies pursuing siloed proof of concepts [26].

7 Summary and conclusion

AI offers companies significant opportunities to reduce inefficiencies, improve outcomes and transform their business models [27]. Responsible AI will help minimise undesirable, unintended consequences, such as reputational damage. As AI permeates greater swathes of private and public sector organisations, managing this risk becomes ever more important.

Principles for responsible AI are an important starting point but will only deliver what companies need if they are combined with governance practices that help to shepherd an AI application from proof of concept to delivery at scale.

This means developing ethical and technical frameworks that can be unambiguously represented in software. These frameworks need to be rigorously tested and measured continuously to ensure the system remains ethical and effective throughout its lifecycle.

The approach will help protect businesses from risk and engender trust in their services from consumers, which in turn will help to drive use. This trust will be important to win as AI is increasingly asked to make important and even life-changing decisions, such as whether an individual

should be given a mortgage, or what medical treatment they should receive.

AI has made rapid progress into enterprise applications over the past few years. If it is to progress further and be deployed on an industrial scale across multiple use cases, governance around ethical principles will be critical.

Copyright © 2020 Accenture. All rights reserved. Accenture, and its logo are trademarks of Accenture.

This document is intended for general informational purposes only and does not take into account the reader's specific circumstances and may not reflect the most current developments. Accenture disclaims, to the fullest extent permitted by applicable law, any and all liability for the accuracy and completeness of the information in this presentation and for any acts or omissions made based on such information. Accenture does not provide legal, regulatory, audit, or tax advice. Readers are responsible for obtaining such advice from their own legal counsel or other licensed professionals.

This document makes reference to marks owned by third parties. All such third-party marks are the property of their respective owners. No sponsorship, endorsement or approval of this content by the owners of such marks is intended, expressed or implied.

8 About Accenture

Accenture is a leading global professional services company, providing a broad range of services in strategy and consulting, interactive, technology and operations, with digital capabilities across all of these services. We combine unmatched experience and specialized capabilities across more than 40 industries—powered by the world's largest network of Advanced Technology and Intelligent Operations centers. With 513,000 people serving clients in more than 120 countries, Accenture brings continuous innovation to help clients improve their performance and create lasting value across their enterprises. Visit us at <http://www.accenture.com>.

Funding Not applicable.

Compliance with ethical standards

Conflict of interest Author works for a company that undertakes business in the deployment of AI systems as part of its commercial activities.

Availability of data and material Not applicable.

Code availability Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. The Institute for Ethical AI & Machine Learning: The responsible machine learning principles. <https://www.ethical.institute/principles.html>. Accessed 27 July 2020
2. UNI Global Union: Top 10 principles for ethical artificial intelligence. https://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf (2017). Accessed 27 July 2020
3. Microsoft: Responsible AI. <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1:primaryr6/>. Accessed 27 July 2020
4. Organisation for Responsible Business: What is responsible business. <https://www.orbuk.org.uk/what-is-responsible-business/>. Accessed 27 July 2020 and Online document Global Initiatives. Responsible Business. <https://www.responsiblebusiness.com/>. Accessed 27 July 2020
5. Gartner: Gartner says nearly half of CIOs are planning to deploy artificial intelligence. <https://www.gartner.com/en/newsroom/press-releases/2018-02-13-gartner-says-nearly-half-of-cios-are-planning-to-deploy-artificial-intelligence> (2018). Accessed 27 July 2020
6. Kanioura, A., Lucini, F.: Ready. Set. Scale. Accenture. https://www.accenture.com/us-en/insights/artificial-intelligence/scaling-enterprise-ai?src=SOMS&c=acn_glb_scalingenterpriexacttarget_t_11099764&n=emc_0120 (2019). Accessed 27 July 2020
7. Accenture: AI ethics and governance. <https://www.accenture.com/us-en/services/applied-intelligence/ai-ethics-governance>. Accessed 27 July 2020
8. Marr, B.: Is artificial intelligence dangerous? 6 AI risks everyone should know about. Forbes. <https://www.forbes.com/sites/bernardmarr/2018/11/19/is-artificial-intelligence-dangerous-6-ai-risks-everyone-should-know-about/#5920f7dc2404> (2018). Accessed 27 July 2020
9. Dastin, J.: Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (2018). Accessed 27 July 2020
10. Hunt, E.: Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter> (2016). Accessed 27 July 2020
11. Gralla, P.: Amazon Prime and the racist algorithms. Computerworld. <https://www.computerworld.com/article/3068622/amazon-prime-and-the-racist-algorithms.html> (2016). Accessed 27 July 2020
12. Brown, T.: The AI skills shortage. IT chronicles. <https://itchronicles.com/artificial-intelligence/the-ai-skills-shortage/> (2019). Accessed 27 July 2020

13. Anem, R.: Software testing skills must shift with AI. <https://www.red-gate.com/simple-talk/opinion/career/software-testing-skills-must-shift-ai/> (2018). Accessed 27 July 2020
14. Lomas, N.: UK Uber drivers are taking the algorithm to court. TechCrunch. <https://www.techcrunch.com/2020/07/20/uk-uber-drivers-are-taking-its-algorithm-to-court/> (2020). Accessed 27 July 2020
15. Gartner (2019) Gartner survey shows 37 percent of organizations have implemented AI in some form. Gartner. <https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have>. Accessed 4 Aug 2020
16. Asimov: I Robot. Gnome Press, New York (1950)
17. IEEE: The IEEE Global Initiative on ethics of autonomous and intelligent systems. https://www.standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_general_principles.pdf. Accessed 4 Aug 2020
18. House of Lords Select Committee on Artificial Intelligence: AI in the UK: ready, willing and able? Parliament publications. <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf> (2018). Accessed 4 Aug 2020
19. EC: Ethics guidelines for trustworthy AI. <https://www.ec.europa.eu/futurium/en/ai-alliance-consultation#:~:text=The%2520Ethics%2520Guidelines%2520for%2520Trustworthy,strategy%2520announced%2520earlier%2520that%2520year> (2019). Accessed 4 Aug 2020
20. Accenture: Responsible AI: governance guidebook, pp. 20–21. <https://view.pagetiger.com/aigovernanceguide> (2019). Accessed 4 Aug 2020
21. Brake: The Road Safety Charity (2020) Driver gender—key facts. <https://www.brake.org.uk/facts-resources/1593-driver-gender>. Accessed 20 July 2020
22. Talemantes, J.: What is red teaming and why do i need it? Red Team. <https://www.redteamsecure.com/blog/what-is-red-teaming-and-why-do-i-need-it-2/>. Accessed 20 July 2020
23. West, S.M., Whittaker, M., Crawford, K.: Discriminating systems: gender, race and power in AI. AI Now Institute. <https://www.ainowinstitute.org/discriminatingystems.pdf> (2019). Accessed 28 July 2020
24. Perez, C.C.: Invisible Women. Vintage Publishing, London (2019)
25. Accenture: Scaling enterprise AI for business value. <https://www.accenture.com/gb-en/insights/artificial-intelligence/scaling-enterprise-ai?src=SOMS> (2019). Accessed 28 July 2020
26. Accenture: AI: built to scale. <https://www.accenture.com/us-en/insights/artificial-intelligence/ai-investments?src=SOMS>. Accessed 20 July 2020
27. Eitel-Porter, R.: AI explained—a guide for executives. Accenture, pp. 22–25. www.Accenture.com/AI-explained (2018). Accessed 20 July 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations